18TH WORLD CONFERENCE ON EARTHQUAKE ENGINEERING
WCEE2024
MILAN

MILAN, ITALY
30TH JUNE - 5TH JULY 2024
www.wcee2024.it

MACHINE LEARNING APPROACH FOR PREDICTING THE FLEXURAL BEHAVIOR OF STEEL MEMBERS

Cyrus Eshaghi [1], Elisa Cerqueira [2], Xavier Romão [2] & José Miguel Castro [2]

[1] Construct, Faculty of Engineering University of Porto, Portugal, cyruses@fe.up.pt

[2] Construct, Faculty of Engineering University of Porto, Portugal

Abstract: *This paper investigates the influence of manufacturing tolerances on the cyclic behavior of steel members with European I and H-shaped profiles. It also explores the applicability of machine learning methods in predicting this behavior. Previous research has focused on nominal sections, overlooking the effects of dimensional tolerances specified in the EN10034 standard. Through advanced finite element modeling, the study evaluates the influence of geometrical variability on the behavior of steel members under cyclic flexural loading. Then, a total set of 2300 samples was generated using Latin Hypercube sampling for a range of profiles (IPE300 to IPE600) and (HEB300-HEB550) with different lengths to evaluate the effect of dimension tolerances on the behavior of beams and columns. The results showed that the variability in member behavior is significant and can cause variation in overstrength ratio, energy dissipation, and rotation capacity. Then, some models were developed using machine learning techniques to predict rotation and moment at key points of backbone curves. To this aim, various models were evaluated, including nonlinear and linear regression analysis, neural network, decision tree, and random forest.*

1 Introduction

The latest international codes encourage the use of advanced numerical modeling and design checks for a thorough seismic assessment of structures. In particular, nonlinear structural analysis employing concentrated plasticity phenomenological models effectively simulates the flexural behavior of steel members that experience cyclic deterioration. However, these models' accuracy and acceptance criteria heavily rely on the rotation and bending strength parameters utilized, which need to be calibrated based on experimental results or advanced finite element models that accurately replicate the physical behavior of the structural component.

Several studies have suggested various parameters for nonlinear modeling that could be utilized to calibrate concentrated plasticity models. For example, Lignos and Krawinkler [1] established empirical equations for estimating plastic rotations prior to and after capping, as well as the rate of cyclic deterioration in special moment connections. Additionally, they provided quantitative information for determining moment capacity at capping and residual moment after cyclic deterioration. Araujo et al. [2] proposed empirical equations for estimating plastic rotations in steel beam-columns with European profiles subjected to monotonic and cyclic loading. Mohabeddine et al. [3] derived empirical equations for estimating rotation capacity in steel beams with European profiles, which can be utilized as acceptance criteria for performance-based assessment of structures or for calibration purposes in push-over analysis. Lignos et al. [4] suggested modeling criteria for the first-cycle envelope and monotonic backbone curves of steel wide-flange columns, which can be employed in nonlinear static and dynamic frame analyses. Previous studies that have examined the cyclic behavior of

steel moment-resisting frame components have primarily been based on nominal geometric dimensions provided by manufacturers. However, actual cross-sections often deviate considerably from the nominal dimensions, as outlined by the tolerance limit deviations specified in the EN 10034:1993 standard [5]. Furthermore, while the standard provides acceptable intervals for section width, height, and thickness, the local slenderness of the plates within those intervals can vary significantly for a given profile. Since local slenderness is a critical factor in the flexural behavior of steel beams subjected to large deformation, understanding the extent of this effect is crucial.

Several researchers have explored the impact of these types of geometrical imperfections on the design and load-carrying capacity of steel members subjected to monotonic loading, including Byfield and Nethercott [6], Melcher et al. [7], and Kala et al. [8]. However, to the authors' knowledge, studies have yet to examine the impact of cross-section dimension variability on the cyclic behavior of steel beams. While existing experimental tests may include these effects, large numbers of tests would be required to identify the impact of these imperfections, making such an approach infeasible therefore, the utilization of advanced finite element simulation can greatly assist in addressing this issue.

The use of machine learning (ML) methods in civil engineering research is a rapidly growing field due to its potential to accurately predict the parameters and behavior of structures while saving time and reducing analysis costs. However, enough accurate data is essential for training ML models since it impacts their performance, generalization, and capabilities. The large data set available in this research can be employed to accurately train a model to predict the nonlinear behavior of steel members. For steel members, several researchers attempted to employ ML models: mainly ANNs [27-30]. Abdalla et al. [27] trained deep ANNs using 11 specimens to predict the moment-rotation response of single web angle and shear-tab connections. The model used only three features. De Lima et al. [28] trained an ANN using 26 specimens to predict the elastic stiffness and plastic strength of bolted extended endplate connections. Ghassemieh et al. [31] developed an ANN model to predict the trilinear response of 8-bolt extended endplate connections with plate rib stiffeners. The model was trained using a total of 25 data point generated by 3-dimensional finite element (FE) simulations; the FE model was validated against two test specimens. Faridmehr et al. [29] trained an ANN model using data from test specimens of connections with top, seat and web angles. The model performed better compared to Eurocode 3 component method with respect to elastic stiffness and plastic strength, where the observed errors were mostly high. Kueh [30] developed an ANN model for predicting the elastic stiffness and ultimate strength of endplate connections. The model was trained using a dataset of 52 physical and FE-simulated specimens. Although these models were found to be of better performance compared to other empirical models, their performance remained limited. This is because a limited amount of data was used in the models' development. This in turn affects the quality of model training and the model's ability to capture the effect of all significant response predictors. This is particularly detrimental for sensitive response parameters such as ductility parameters. Ning et al. [8] explored the potential of using three deep learning (DL) models to predict the nonlinear time-history responses of civil engineering structures under seismic loading. Afshari et al. [9] reviewed DL-based methods in structural reliability analysis (SRA). They found that while ML-based techniques can improve SRA accuracy, their performance is limited for high-dimensional and nonlinear problems. Mylvaganam et al. [10] conducted a systematic review of concrete strength prediction models. They concluded that while ML models can estimate the strength of novel concrete with superior accuracy, the black-box nature of the process is a disadvantage. Luo et al. [26] presented a machine learning-based model (ML-BCV) that rapidly predicts backbone curves for flexure- and shear-critical concrete columns, outperforming traditional modeling approaches by significantly reducing root-mean-square error and demonstrating increased robustness and accuracy.

This study involved 1600 finite element simulations in ABAQUS of cantilever beams (IPE300-IPE600) and 600 models of HEB300 to HEP450 with and without axial loads subjected to cyclic loading to investigate the effect of manufacturing tolerances on the rotation capacity of beams.

The results of FE modeling were used to develop different types of ML models to predict variability in nonlinear modeling parameters of I and H-shaped profiles subjected to cyclic loading, considering manufacturing tolerances. By utilizing ML techniques, this research aims to improve the accuracy of predicting the behavior of structures and ultimately enhance their design and safety.

2 Methodology

To accomplish the goals, a high-fidelity finite element model is created using ABAQUS software. To model the cyclic behavior, standardized SAC loading protocol is adopted [17]. The model is a cantilever beam, representing the beam's behavior from one end to the inflection point. This approach is commonly used in experimental and numerical studies. The study focuses on IPE300 to IPE600 profiles, which are widely used in practice [11,12], and considers a member with a length of 2, 2.5, 3, 3.5m and also HEB300 to HEB450 profiles with lengths equal to 2m without and with 30% axial force. Figure 1 shows a 3D schematic view of the model with one edge fully restrained in all 6 degrees of freedom to ensure fixed boundary conditions. The cyclic loading is applied on the other edge, which is restrained in the X-X direction to prevent out-of-plane displacements at the top of the specimen.

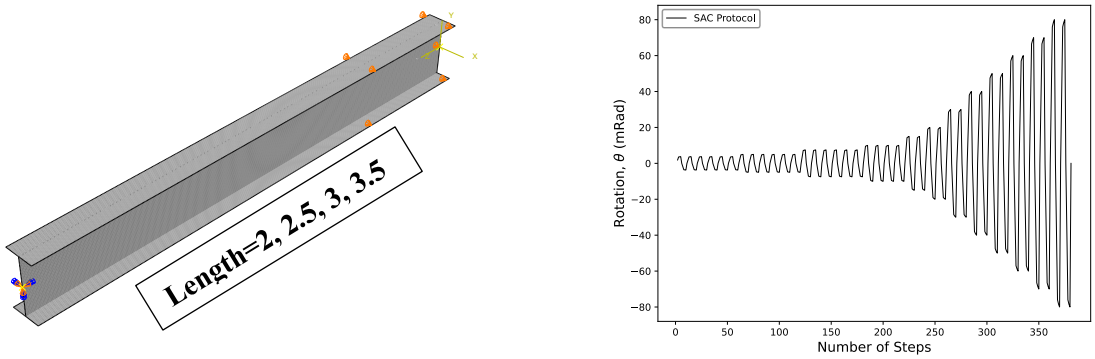


Figure 1 Model of beams in Abaqus and SAC protocol

To prevent Lateral-Torsional Buckling (LTB), lateral restrictions were added to the flanges. According to ANSI/AISC 341-16 [13], the unbraced length (L_b), which is the distance between the lateral restraints, was calculated using Eq. 1

$$L_b = 0.095 i_z \frac{E}{F_y} \quad (1)$$

Where i_z is the radius of gyration of the cross-section, E is the young modulus, F_y is the yield stress obtained from coupon tests.

The constitutive material model parameters for the Voce-Chaboche model, which combines kinematic and isotropic models, were obtained using the following equation.

$$\sigma^0 = \sigma|_0 + Q_\infty(1 - e^{-be^{pl}}) \quad (2)$$

In Equation (2), the term σ^0 represents the change in the yield surface size as a function of equivalent plastic strain e^{pl}. The yield stress at zero equivalent plastic strain is represented by $\sigma|_0$, while Q_∞ indicates the maximum change in the yield surface size. Parameter b describes the rate at which the yield surface size changes as plastic strain increases. The nonlinear kinematic hardening component involves the movement of the yield surface in the stress space through the vector backstress α, which is used to capture complex phenomena like the Bauschinger effect. Equation (3) represents the backstress evolution law, a mathematical expression of the vector function that shifts the center of the Mises yield surface.

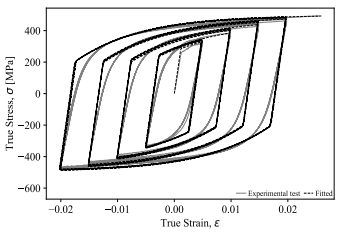
$$\dot{\alpha_k} = C_k \frac{1}{s^0}(\sigma - \alpha)\dot{e}^{pl} - \gamma_k \alpha_k \dot{e}^{pl} \quad (3)$$

where C_k and γ_k are material parameters calibrated from stabilized cycle, σ is the stress matrix, $\dot{\alpha_k}$ is the evolution of the back stress, and $\dot{e}^{pl}$ is the equivalent plastic strain rate. The subscript $''_k''$ defines the backstress number. The constitutive parameters for this model were obtained through an optimization process using data from Chen et al.'s coupon test [14]. A comparison between the coupon test and the Voce-Chaboche model is illustrated in Figure 2, and Table 1 lists the parameters used to describe the constitutive model. The

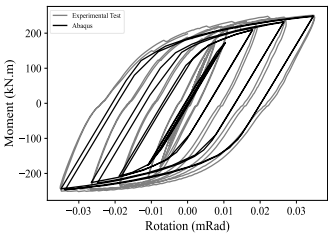
finite element model was then validated against various experimental tests on different members, and the outcomes are depicted in Figure 2.

Table 1 Nonlinear Isotropic and Kinematic hardening parameters.

Material	σ|0 (N/mm2)	Q∞ (N/mm2)	b	C (N/mm2)	γ
Q355b/S355	285	69	2.86	14238	96.15

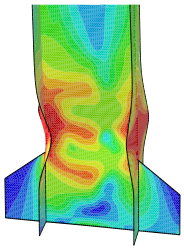


Experimental Coupon from Chen et al. [14]



Experimental Test on IPE300 D'Aniello et al. [15]

Experimental Test on IPE300 D'Aniello et al. [15]



Abaqus Simulation of IPE300

Figure 2 Model Validation

3 Dimension Characteristics Variability

Table 2 presents the tolerances on the shape dimensions of I and H sections of structural steel, as specified by the European standard [5]. The descriptions of the profile dimensions are illustrated in Figure 3.

Table 2-Dimensional tolerance for structural steel I and H sections (EN 10034:1993) (units in mm)

Section Height (h)		Flange Width (b)		Web Thickness (tw)		Flange Thickness (tf)	
h<=180	+3 -2	b<=110	+4 -1	t_w<7	±0.7	t_f<6.5	+1.5 -0.5
180<h<=400	+4 -2	110<b<=210	+4 -2	7=< t_w <10	±1	6.5=<t_f<10	+2 -1
400<h<=700	+5 -3	210<b<=325	+4 -4	10=< t_w <20	±1.5	10=<t_f<20	+2.5 -1.5
h>700	+5 -5	b>325	+6 -5	20=< t_w <40	±2	20=<t_f<30	+2.5 -2
				40=< t_w <60	±2.5	30=< t_f <40	+2.5 -2.5
				60< t_w	±3	40=< t_f <60	+3 -3
						60< t_f	+4 -4

Table 2 shows that the code allows for significant changes in thicknesses, which can greatly affect the slenderness ratios and overall performance of the profile. The coefficient of variation for geometric imperfections is based on research from the 1970s, but some studies have questioned whether these estimates are still valid due to advancements in manufacturing methods [6]. Melcher et al. [7,16] conducted experimental research on Czech steel hot-rolled IPE profiles, and statistically evaluated the geometrical characteristics of the cross-section dimensions, as presented in Table 3 and Table 4. This study used the results from those tables to define the next steps in the paper.

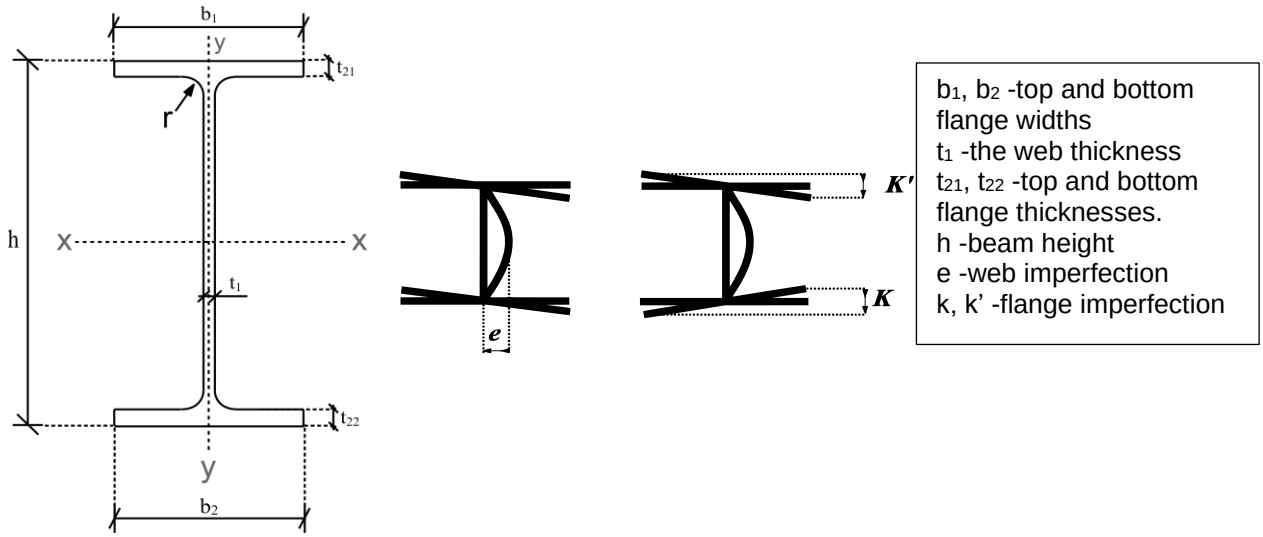


Figure 3 Cross section geometry and imperfections

As seen in Fig. 3, h is the height of the beam, b_1, b_2 are the top and bottom flange widths, t_1 is the web thickness, and t_{21}, t_{22} are the top and bottom flange thicknesses, The web imperfection in the shape of the profile, denoted as e, can be expressed as 0.004h, where h represents the height of the profile. The flange imperfection in shape, denoted as k + k', can be expressed as 0.02b, where b represents the width of the flange.

Table 3 Relative statistical geometric characteristics [7]

Quantity	Mean	Standard deviation	Skewness	Kurtosis	Min Value	Max value
h	1.001	0.00443	-0.4063	3.0150	0.989	1.013
b_1	1.012	0.01026	-0.3939	4.239	0.975	1.049
b_2	1.015	0.00961	-0.5448	3.887	0.975	1.037
t_1	1.055	0.04182	1.0545	7.4730	0.949	1.3
t_{21}	0.988	0.04357	-0.2991	2.663	0.880	1.094
t_{22}	0.998	0.04803	0.3303	2.766	0.858	1.129

Table 4 Correlation matrix of geometric characteristics [7]

Quantity	h	b_1	b_2	t_1	t_{21}	t_{22}
h	1	-0.0068	0.0534	0.0399	-0.0686	-0.0989
b_1	-0.0068	1	0.6227	-0.2142	-0.2681	-0.1456
b_2	0.0534	0.6227	1	-0.2132	-0.1596	-0.0423
t_1	0.0399	-0.2142	-0.2132	1	0.2368	0.2451
t_{21}	0.0686	-0.2681	-0.1596	0.2368	1	0.7634
t_{22}	-0.0989	-0.1456	0.0423	0.2451	0.7634	1

4 Stochastic Model

Latin hypercube sampling (LHS) method was used to define the random geometries for finite element analysis. The LHS method is computationally efficient compared to standard Monte Carlo sampling since it reduces the number of samples significantly [19-20]. However, the results of LHS sections are dependent on whether the components of X (variables) are independent or dependent. In this paper, a procedure for producing Latin hypercube samples is used based on a correlation matrix of geometric characteristics, as presented in Table 4, since section variables are dependent. An experimental distribution is defined (after analyzing the data, it was determined that they do not follow any specific distribution) to capture the true distribution of the variables.

For each IPE profile considered in this study, 50 samples with varying dimensions are generated using the LHS method and this process repeated for 4 different lengths of each profile. Figure 4 displays four of the six input variables as dimensional tolerances that were generated using Latin hypercube sampling (LHS).

After this stage, dimensions that do not match Tables 2 and 3 will be removed. Each profile has four series of these dimensions with lengths equal to 2, 2.5, 3, and 3.5 meters and 50 possible combinations of 6 random variables in cross-section dimensions. Lastly, ABAQUS ran 200 models for each of the eight profiles, and in total, 1600 models of all profiles. Along with the six random variables mentioned earlier, all models also included a constant imperfection in the flange and web based on EN 10034:1993. This imperfection equals 0.02b ('b' represents the average width of the top and bottom flanges) in flanges, and 0.004h (h is the height of the beam) in the web. 50% of these imperfections were included in all models as constant because, during sensitivity analysis, it was found that they have a minimal impact on the results. Therefore, they are considered constant for each profile, based on its dimensions and not random variables.

5 Sensitivity Analysis

A sensitivity analysis was conducted to determine the separate impact of each random variable. In one analysis, $b_{1,2}$ was considered a random variable while the other dimensions were nominal, namely h, $t_{21,22}$, and t_1, respectively. As Figure 4 illustrates, the effect of b and h is relatively minor. This is due to the low dimension tolerances specified in EN10034 for b and h, compared to their actual dimensions. For example, For IPE 300, the maximum tolerances for h and b are (-0.66%, +1.33%) and (-1.33%, +2.67%), respectively. In terms of t_2 (flange thickness) and t_1(web thickness), the tolerances are (-14%, +23.36%) and (-14.08%, +14.08%), respectively. However, since changes in the thickness of the top and bottom flanges were considered together, one of them can affect the other and reduce the impact of t_2 compared to t_1.

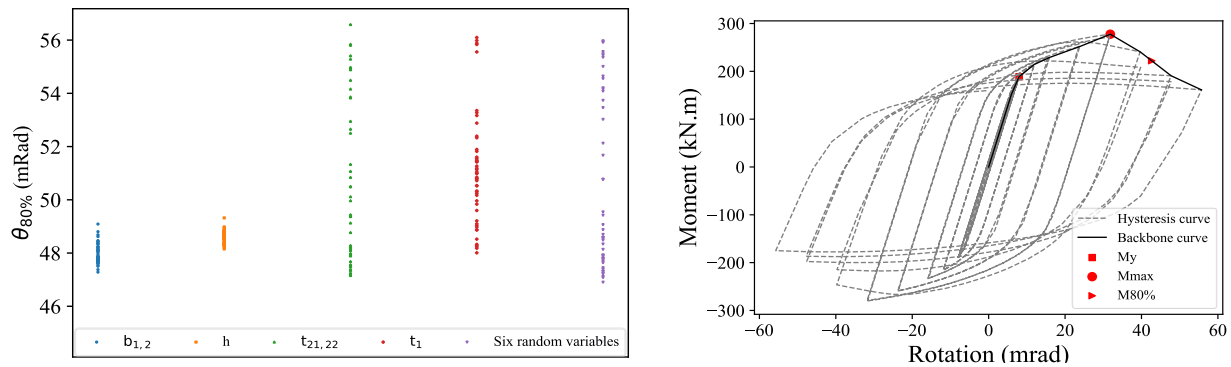


Figure 4 Sensitivity of $\theta_{80\%}$ to each individual random variable (left), hysteresis moment-rotation curves (right)

6 Results and Discussion

Figure 5 displays the cyclic moment-rotation curve and the first cycle envelope which illustrating the behavior of one of the 50 models of IPE300 profile with length equal to 2m, which is determined by moment and rotation values at yielding, maximum moment (θ_{max}, M_{max}), and an 20% reduction in strength ($\theta_{80\%}$, $M_{80\%}$). Figure 5 shows the first cycle curves for IPE300 to IPE600, based on 1600 samples for the profiles with length equal to 2, 2.5, 3, 3.5m. The results indicate that changes in the dimensions of the beams significantly affect their strength and rotation capacity. The average first cycle curve for each profile (green line) has higher rotation and strength than the nominal cross section curves. However, higher strength can increase the over-strength factor, which is not desirable for seismic design, and lower rotation capacity can decrease the seismic performance of the structure. These results could be useful for reliability analysis and risk assessment, considering the variability of input modeling parameters.

The rotation at strength drops $\theta_{80\%}$ is widely accepted as the parameter that defines the rotation capacity of the beam when subjected to large deformation.

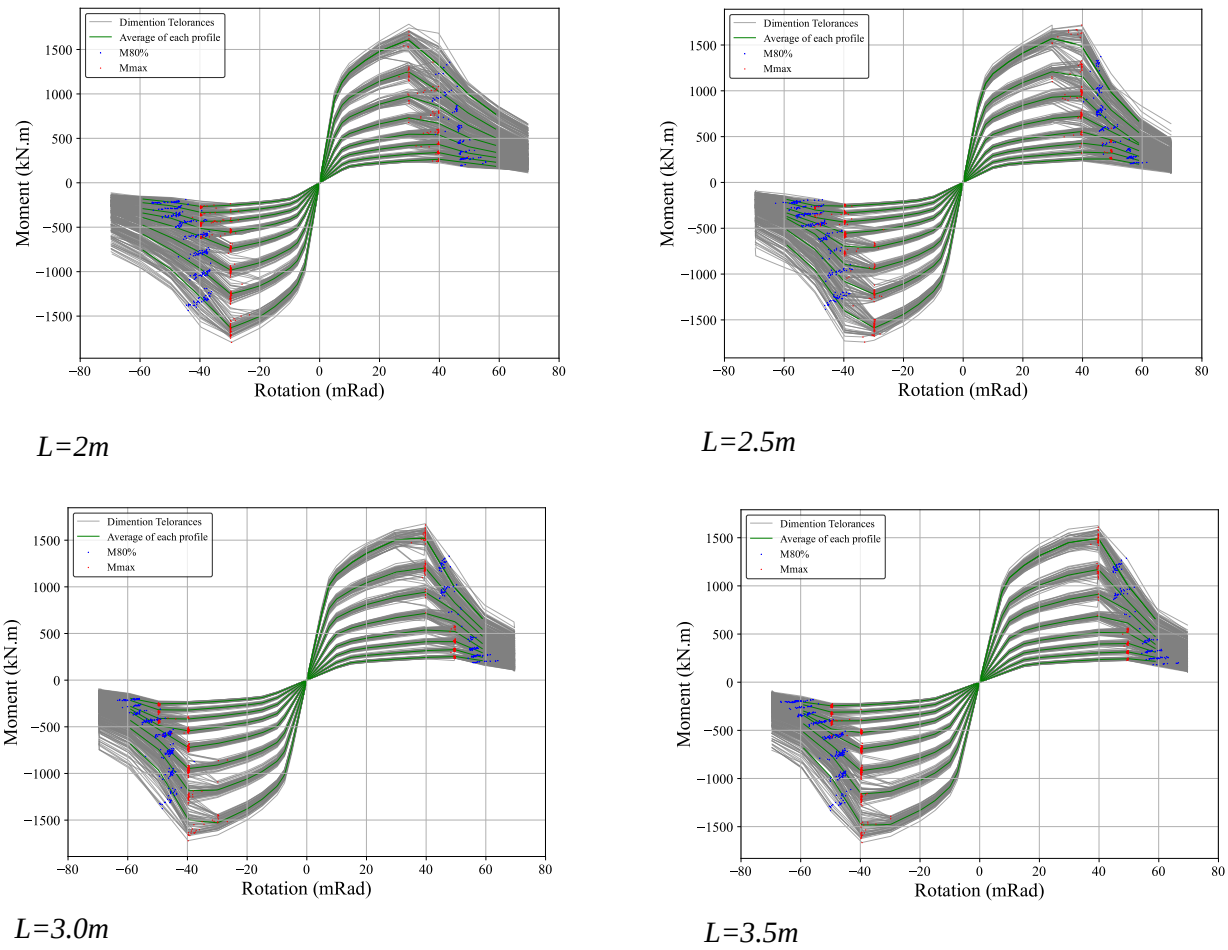
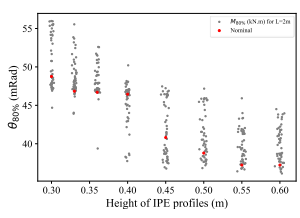
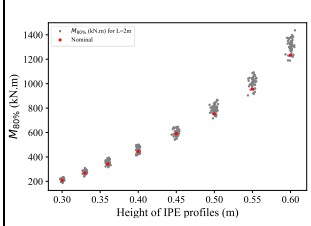


Figure 5 Variation in Backbone Curves of IPE300 to IPE600, L=2, 2.5, 3, 3.5m

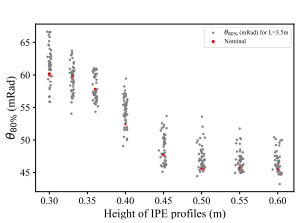
Figures 6 also illustrates variability of the (θ_{max}, M_{max}) and ($\theta_{80\%}$, $M_{80\%}$) values for various profiles with a length of 2 meters. The manufacturing tolerances in the dimensions of the cross-section have a significant impact on the variability of the parameters. The range of variation differs across different profiles, as depicted in Figure 6. The reason is that EN 10034:1993 defines a range for dimension tolerances for a series of profiles, and this can increase the slenderness of a profile in the beginning of the range more than the profile at the end of the range.



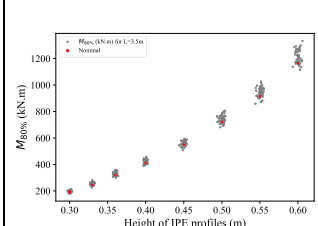
a



b



c



d

Figure 6 (a) Variability of $\theta_{80\%}$, L=2m and (b) variation in $M_{80\%}$, L=2m. (c) Variability of $\theta_{80\%}$, L=3.5m and (d)` variation in $M_{80\%}$, L=3.5m.

7 Predictive Models

The following section involved creating several models using machine learning methods to predict the maximum moment and rotation capacity and evaluating the accuracy of these models. Moreover, a numerical equation was developed to predict the rotation capacity at maximum moment and when the maximum moment has decreased by 20%. Several models were evaluated, including nonlinear regression, linear regression analysis, neural network, decision tree, and random forest, with the aim of achieving the desired outcome. To evaluate the models, 80% of the data was used for training, while the remaining 20% was used for testing. Additionally, the training data was divided into two sets: a training set and a validation set, which were used to optimize the hyperparameters of the models under consideration. The models were created individually for both IPE and HEB profiles, with a particular emphasis on IPE profiles due to the larger amount of data generated during this research specifically for this type of profile.

7.1 Nonlinear Regression Analysis

Careful consideration of model assumptions is crucial in nonlinear regression analysis, including the distribution of errors, as the choice of model and parameters can significantly impact the results. To ensure the model's appropriateness and the reliability of results, it is important to evaluate the model's fit to the data using diagnostic plots or statistical tests. Current construction steels can undergo substantial deformation without loss of strength, and the rotation capacity of a steel member is heavily dependent on geometrical instabilities. Flange and web slenderness are the major influencing parameters of a given plated beam. Thus, an analytical equation, which employs nonlinear regression analysis to estimate rotations corresponding to a 20% drop in the maximum moment for IPE profiles (range between IPE300-IPE600), and HEB profiles with 30% axial force is presented in Equation 5.

$$\theta_{80\%}(mRad) = a\left(\frac{L}{h}\right)^{b}\left(\frac{h_w}{t_w}\right)^{c}\left(\frac{b_f}{2t_f}\right)^{d} \quad (5)$$

Where $\frac{h}{t_w}$ represent the web slenderness and $\frac{b}{2t_f}$ represent flange slenderness. The following are the regression coefficients:

Table 5 Coefficient of equation (5) for IPE profiles

Dataset	a	b	c	d
IPE	273.35	0.16	-0.45	-0.38

Dataset	a	b	c	d
HEB	387	0.26	-0.61	-0.52

L *Length of cantilever beam*
h *Beam depth*
t_f *Average of flange thickness*
b_f *Average of flange widths*
t_w *Web thickness*
h_w *Web height*

To measure the accuracy of the models three metrics including mean absolute error (MAE), mean square error (MSE) and R^2 were used. There is a strong correlation between measured and anticipated results, as illustrated in figure 7 and Table 6.

$$MAE = \frac{1}{n}\sum\left|y_{pred} - y_{true}\right| \quad (6)$$

$$MSE = \frac{1}{n}\sum(y_{pred} - y_{true})^2 \quad (7)$$

n *Number of samples in the dataset*
y_{pred} *Predicted value of the model*
y_{true} *True value of the target variable*

Table 6 Accuracy of nonlinear regression analysis

Prediction IPE	MSE	MAE	R^2
θ_{max}	1.9	6.4	0.85
$\theta_{80\%}$	3.35	1.46	0.94

Prediction HEB	MSE	MAE	R^2
θ_{max}	9.19	2.46	0.54
$\theta_{80\%}$	1.2	0.94	0.73

The proposed equation underwent extensive research before being suggested as indicated in table 5, The

benefit of this approach is that we arrive at a final equation that researchers and engineers may utilize with ease.

7.2 Linear Regression

The linear regression method is widely used in civil engineering to model the relationship between various factors affecting the performance of structures, such as load-bearing capacity, deflection, and durability. In this part a liners regression model developed to predict $\theta_{80\%}$, θ_{max}, M_{max} or $M_{80\%}$. The findings are presented in Table 7.

Table 7 Accuracy of linear regression analysis

Prediction IPE	MSE	MAE	R^2
θ_{max}	7.3	2.1	0.81
$\theta_{80\%}$	4.85	1.78	0.91
$M_{80\%}$	1698.13	33.64	0.9898

Prediction HEB	MSE	MAE	R^2
θ_{max}	5.68	1.89	0.44
$\theta_{80\%}$	0.83	0.72	0.738
$M_{80\%}$	376	15.99	0.993

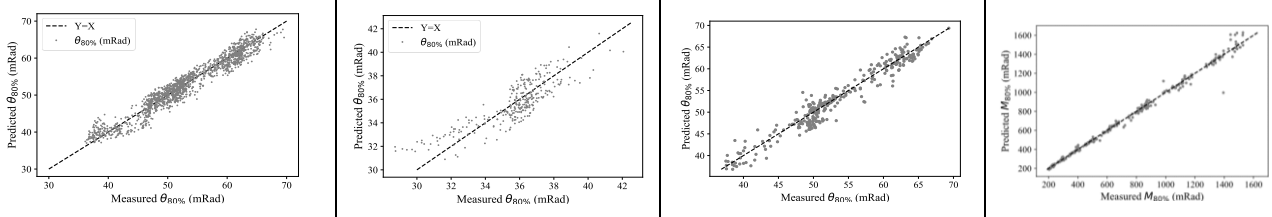
7.3 Neural Networks

Neural networks are a type of machine learning model that consists of layers of interconnected nodes, also known as neurons, which process and transmit information using mathematical functions (figure 7).

Table 8 Input variables

Input Variables	Description
L	Length of cantilever beam
h	Height of profiles
b_1	Top flange width
b_2	Bottom flange width
t_1	Web thickness
t_{21}	Top flange thickness
t_{22}	Bottom flange thickness
L_b	Unbraced length
C	Web height

This study employed 9 input variables (Table 8) to train the models using 1280 samples (80% of the data from IPE profiles), with an additional 320 samples used to test the models' accuracy. Figure 7 displays the prediction of $\theta_{80\%}$ and $M_{80\%}$, while Table 9 presents the accuracy of the models.



IPE profiles- NRA	*HEB profiles- NRA*	*IPE profiles- ANNs*	*HEB profiles- ANNs*

Figure 7 Correlation between measured and anticipated results

Table 9 Accuracy of neural network model

Prediction IPE	MSE	MAE	R^2
θ_{max}	3.03	0.98	0.928
$\theta_{80\%}$	1.53	0.96	0.971
$M_{80\%}$	38.2	3.8	0.9997

Prediction HEB	MSE	MAE	R^2
θ_{max}	4.8	1.27	0.71
$\theta_{80\%}$	1.13	0.83	0.66
$M_{80\%}$	197	10	0.996

7.4 Decision Tree

Decision trees are a versatile tool that can be used for both classification and regression problems. When applied to regression analysis, decision trees aim to predict continuous target variables by dividing the feature space into smaller subsets based on input feature values [22]. Table 10 presents the accuracy of the models.

Table 10 Accuracy of Decision Tree model

Prediction (IPE)	MSE	MAE	R^2
θ_{max}	4.01	0.648	0.907
$\theta_{80\%}$	2.48	0.87	0.956
$M_{80\%}$	171.43	7.44	0.9989

Prediction (HEB)	MSE	MAE	R^2
θ_{max}	9.05	1.4	0.405
$\theta_{80\%}$	1.9	1.03	0.38
$M_{80\%}$	531	17.87	0.990

7.5 Random Forest

One of the key advantages of using random forest for regression is its ability to handle high-dimensional data and outliers, as well as its capability to capture complex nonlinear relationships between the input features and the target variable [23]. Moreover, the significance of the input features can be readily interpreted by analyzing their contributions to the variance reduction in the individual decision trees. Table 11 shows the accuracy and table 12 shows the importance of features in the training set of the model.

Table 11 Accuracy of Random Forest

Prediction (IPE)	MSE	MAE	R^2
θ_{max}	3.35	0.768	0.917
$\theta_{80\%}$	1.65	0.85	0.969
$M_{80\%}$	128.68	7.4	0.992

Prediction (HEB)	MSE	MAE	R^2
θ_{max}	4.48	1.35	0.59
$\theta_{80\%}$	0.72	0.64	0.68
$M_{80\%}$	736	20.16	0.986

Table 12 Importance of Features Random Forest

Features	θ_{max}	$\theta_{80\%}$	$M_{80\%}$
L	0.456	0.492	0.0026
h	0.311	0.213	0.107
b_1	0.069	0.095	0.394
b_2	0.069	0.021	0.224
t_{21}	0.037	0.018	0.0015
t_{22}	0.007	0.018	0.006
t_1	0.012	0.013	0.013
L_b	0.038	0.099	0.207
r	0.0012	0.031	0.048

Conclusion

The primary focus of this study was to examine how dimensional tolerances impact the flexural behaviour of steel beams and columns when subjected to cyclic loading. To accomplish this objective, the study involved 1600 analyses on eight distinct steel IPE profiles, ranging from 300 to 600, and 600 analyses of 6 HEB steel profiles ranging from 300 to 450. After evaluating the results, the study found that this parameter has a significant effect and can potentially cause variation in overstrength ratio, rotation capacity, classification, and energy dissipation of steel beams in MRF and additionally for columns caused variation in axial shortening. Therefore, future design and analysis of steel structures should give greater consideration to this geometrical imperfection. In this study, various ML techniques such as linear and nonlinear regression analysis, neural network, decision tree, and random forest were employed to predict (θ_{max}, M_{max}), and a 20% reduction in strength (θ80%, M80%). The findings suggest that neural network models are more accurate in predicting both variables, making it a reliable approach for predicting moment and rotation. It is interesting to note that utilizing

a neural network can be highly beneficial in this context since running a model with Abaqus can be both time-consuming and expensive.

A large amount of accurate data is essential for the success of machine learning techniques, and the lack of such data is a major challenge. However, by incorporating dimension tolerances across a range of profiles, this challenge can be overcome. Using only nominal sections can result in low-accuracy models while incorporating dimension tolerances allows for more realistic modeling. This, in turn, leads to more accurate results and better machine learning outcomes.

Acknowledgments

"This work is a result of Agenda “R2UTechnologies - modular systems”, contract C644876810-00000019, investment project 48, funded by the Recovery and Resilience Plan (PRR) and by the European Union - NextGeneration EU.", of which the first author is a grantee. This work is also integrated in the R&D activities of the CONSTRUCT Institute on Structures and Constructions (Instituto de I&D em Estruturas e Construções), financially supported by Base Funding UIDB/04708/2020 through national funds of FCT/MCTES (PIDDAC).

References

[1] Lignos, D.G.; Krawinkler, H.; (2011) *Deterioration Modeling of Steel Components in Support of Collapse Prediction of Steel Moment Frames under Earthquake Loading. Journal of Structural Engineering*. 137(11): p. 1291-1302.

[2] Araújo, M.; Macedo, L.; Castro, J.M.; (2017) *Evaluation of the rotation capacity limits of steel members defined in EC8-3*. Journal of Constructional Steel Research. 135: p. 11-29.

[3] Mohabeddine, A.; Koudri, Y.W.; Correia, J.A.F.O.; Castro, J.M.; (2021) *Rotation capacity of steel members for the seismic assessment of steel buildings*. Engineering Structures. 244: p. 112760.

[4] Lignos, D.G.H.; Elkady, A.; Deierlein, G.G.; (2019) *Proposed Updates to the ASCE 41 Nonlinear Modeling Parameters for Wide-Flange Steel Columns in Support of Performance-Based Seismic Engineering*. Journal of Structural Engineering. 145(9): p. 04019083.

[5] Standards, B., BS EN 10034:1993 *Structural steel I and H sections Tolerances on shape and dimensions*. (1993).

[6] Byfield, M. P.; Nethercot, D A.; (1997) *Material and geometric properties of structural steel for use in design*. The Structural Engineer. 75(25).

[7] Melcher, J.K., Holicky´, M. Fajkus, M. Rozlı ´vka, L. (2004) *Design characteristics of structural steels based on statistical analysis of metallurgical products*. Journal of Constructional Steel Research. 60(3-5): p. 795-808.

[8] Ning, C., Xie, Y., & Sun, L. (2023). *LSTM, WaveNet, and 2D CNN for nonlinear time history prediction of seismic responses*. Engineering Structures, 286, 116083.

[9] Afshari, S. S., Zhao, C., Zhuang, X., & Liang, X. (2023). *Deep learning-based methods in structural reliability analysis: a review. Measurement Science and Technology*.

[10] Mylvaganam, N., & Elakneswaran, Y. (2023). *A Systematic Review and Assessment of Concrete Strength Prediction Models*. Case Studies in Construction Materials, e01830.

[11] Mohabeddine, A. I., Eshaghi, C., Barros, R. C. D., Correia, J. A., & Castro, J. M. (2022). *Variability in the Nonlinear Modelling Parameters of Steel I‐shaped Beams Subjected to Cyclic Loading Considering Manufacturing Tolerances*. ce/papers, 5(4), 748-754.

[12] Mohabeddine, A.I.; Eshaghi, C.; Correia, José A.F.O.; Castro J. M. (2022) *Analytical investigation of the cyclic behaviour of I-shaped steel beam with reinforced web using bonded CFRP*. Steel and Composit Structures. 43(4): p. 447-456.

[13] ANSI/AISC. (2016). *Seismic Provisions for Structural Steel Buildings*. In AMERICAN I. Chicago.

[14] Chen, Y.; Wei, S.;Tak-Ming, C. (2014) *Cyclic stress-strain behavior of structural steel with yieldstrength up to 460 N/mm²*. Frontiers of Structural and Civil Engineering. 8(2): p. 178-186.

[15] D'Aniello, M..; Landolfo, R..; Piluso, V.Rizzano, G.; (2012) *Ultimate behavior of steel beams under non-uniform bending. Journal of Constructional Steel Research*. 78: p. 144-158.

[16] Kala, Z.; Kala, J.; (2005) *Sensitivity analysis of lateral buckling stability problems of hot-rolled steel beams*. Slovac Journal of Civil Engineering. 2: p. 9-14.

[17] Mohabeddine, A. I.; Kuidri, Y.W.; Castro, J. M.; Correia, J. A. F. O. (2018) *Numerical Simulation and Calibration of the Cyclic Behavior of Structural Steel Under Different Loading Protocols*, in XIX International Colloquium on Mechanical Fatigue of Metals. Porto, Portugal. p. 97–98.

[18] CEN, Eurocode 3: Design of steel structures-part 1.1: *General rules and rules for buildings, in CEN, European Committee for Standardization*. (1992).

[19] Tang, B.; (1988) *Orthogonal Array-Based Latin Hypercubes. Journal of the American Statistical Association*. 88: p. 1392–1397.

[20] Anders M. J. Olsson, G.E.S.; (2002*) Latin Hypercube Sampling for Stochastic Finite Element Analysis. Journal of Engineering Mechanics*. 128(1).

[21] Stein, M.; (1987) *Large Sample Properties of Simulations Using Latin Hypercube Sampling.* Technometrics. 29(2): p. 143-151.

[22] Witten, I. H.; Frank, E.; Hall, M. A. (2011) *Data Mining: Practical Machine Learning Tools and Techniques* (3rd edition).

[23] Breiman, L. (2001). *Random forests. Machine learning*. 45(1): p. 5-32.

[24] Cheng, X.; Chen, Y.; (2018) *Ultimate strength of H-sections under combined compression and uniaxial bending considering plate interaction*. Journal of Constructional Steel Research. 143: p. 196-207.

[25] Elkady, A., Güell, G., & Lignos, D. G. (2020). *Proposed methodology for building-specific earthquake loss assessment including column residual axial shortening*. Earthquake Engineering & Structural Dynamics, 49(4), 339-355.

[26] Luo, H., & Paal, S. G. (2018). *Machine learning–based backbone curve model of reinforced concrete columns subjected to cyclic loading reversals*. Journal of Computing in Civil Engineering, 32(5), 04018042.

[27] Abdalla, K. M., and Stavroulakis, G. E. (1995). "*A backpropagation neural network model for semi-rigid* 542 *steel connections*." Computer-Aided Civil and Infrastructure Engineering, 10(2), 77-87, DOI: 543 10.1111/j.1467-8667.1995.tb00271.x.

[28] De Lima, L. R. O., Vellasco, P. C. G. d. S., De Andrade, S. A. L., Da Silva, J. G. S., and Vellasco, M. M. 568 B. R. (2005). "*Neural networks assessment of beam-to-column joints*." Journal of the Brazilian 569 Society of Mechanical Sciences and Engineering, 27(3), 314-324, DOI: 10.1590/S1678- 570 58782005000300015.

[29] Faridmehr, I., Nikoo, M., Pucinotti, R., and Bedon, C. (2021). "*Application of component-based* 596 *mechanical models and artificial intelligence to bolted beam-to-column connections*." Applied 597 Sciences, 11(5), 2297, DOI: 10.3390/app11052297.

[30] Kueh, A. B. H. (2021). "*Artificial neural network and regressed beam-column connection explicit mathematical moment-rotation expressions*." Journal of Building Engineering, 43, 103195, DOI: 10.1016/j.jobe.2021.103195.

[31] Ghassemieh, M., and Nasseri, M. (2012). "*Evaluation of stiffened end-plate moment connection through optimized artificial neural network*." Journal of Software Engineering and Applications, 5(3), DOI: 10.4236/jsea.2012.53023